ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois 60439

GLOBAL METHODS FOR NONLINEAR COMPLEMENTARITY PROBLEMS

Jorge J. Moré

Mathematics and Computer Science Division

Preprint MCS-P429-0494

April 1994

Work supported by the Office of Scientific Computing, U.S. Department of Energy, under Contract W-31-109-Eng-38.

ABSTRACT

Global methods for nonlinear complementarity problems formulate the problem as a system of nonsmooth nonlinear equations approach, or use continuation to trace a path defined by a smooth system of nonlinear equations. We formulate the nonlinear complementarity problem as a bound-constrained nonlinear least squares problem. Algorithms based on this formulation are applicable to general nonlinear complementarity problems, can be started from any nonnegative starting point, and each iteration only requires the solution of systems of linear equations. Convergence to a solution of the nonlinear complementarity problem is guaranteed under reasonable regularity assumptions. The converge rate is Q-linear, Q-superlinear, or Q-quadratic, depending on the tolerances used to solve the subproblems.

Global Methods for Nonlinear Complementarity Problems

Jorge J. Moré

1 Introduction

The solution of economic equilibria has been an important motivation for the development of algorithms for nonlinear complementarity problems. The work of Ahn [1] on the PIES (Project Independence Evaluation System) project, in particular, created much interest in the formulation and solution of economic equilibrium problems as complementarity problems. In this paper we are concerned only with the complementarity formulation. For additional information on economic equilibrium problems see the collection of papers edited by Manne [30] and Nagurney [37]; Dirkse and Ferris [10] provide an interesting collection of nonlinear complementarity problems.

The aim of current research on the numerical solution of nonlinear complementarity problems has been to obtain algorithms with global convergence properties. This goal has proved to be elusive. Global methods for the nonlinear complementarity problem

$$x \geq 0, \qquad f(x) \geq 0, \qquad x^T f(x) = 0, \tag{1.1}$$

specified by a mapping $f : \mathbb{R}^n \mapsto \mathbb{R}^n$, either transform (1.1) into a system of nonsmooth nonlinear equations or use continuation to trace a path that leads, under suitable conditions, to a solution of the nonlinear complementarity problem.

In the nonsmooth nonlinear equations approach, the nonlinear complementarity problem is transformed into a system of nonlinear equations $h(x) = 0$ with a mapping $h : \mathbb{R}^n \mapsto \mathbb{R}^n$ that is continuous, but not differentiable everywhere. For example, in the approach studied by Robinson [41], the nonlinear complementarity problem is formulated as the nonsmooth system of nonlinear equations

$$f(x_+) + x_- = 0, \tag{1.2}$$

where $x_+ = \max(x, 0)$ and $x_- = \min(x, 0)$. A computation shows that a solution x^* of (1.2) yields a solution x_+^* of the nonlinear complementarity problem (1.1). Conversely, a solution x^* of (1.1) yields a solution $x^* - f(x^*)$ of the nonsmooth system (1.2). This approach has been pursued, in particular, by Ralph [40], Dirkse and Ferris [9], and Xiao and Harker [47, 48].

Pang [38] has also followed a nonsmooth approach. In Pang's approach the nonlinear complementarity problem is formulated as the constrained system of nonlinear equations

$$\min(x, f(x)) = 0, \qquad x \geq 0. \tag{1.3}$$

Work supported by the Office of Scientific Computing, U.S. Department of Energy, under Contract W-31-109-Eng-38.

Clearly, x^* solves (1.3) if and only if x^* is a solution of the nonlinear complementarity problem. This approach has been developed by Harker and Xiao [19], Pang and Gabriel [39], Gabriel and Pang [14, 15], and Monteiro, Pang, and Wang [32].

The nonlinear complementarity problem (1.1) can also be formulated as the constrained system of nonlinear equations

$$h(x, y) = 0, \qquad x \geq 0, \qquad y \geq 0, \tag{1.4}$$

where $h : \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ is defined by

$$h(x, y) = \begin{pmatrix} f(x) - y \\ Yx \end{pmatrix}, \tag{1.5}$$

and Y is the diagonal matrix $\mathrm{diag}(y_i)$. The nonlinear complementarity problem (1.1) is equivalent to the system (1.4) because if the vector x^* solves problem (1.1), then $h(x^*, y^*) = 0$ for $y^* = f(x^*) \geq 0$, and conversely, if the pair (x^*, y^*) solves (1.4), then x^* solves (1.1) and $y^* = f(x^*)$.

In the continuation approach proposed by McLinden [31] and Kojima, Mizuno, and Noma [24], the nonlinear complementarity problem is transformed into the constrained system of nonlinear equations

$$h(x(\tau), y(\tau)) = \tau \begin{pmatrix} a \\ b \end{pmatrix}, \qquad \tau > 0, \quad (x(\tau), y(\tau)) > 0.$$

Convergence results for this approach show that under suitable conditions there is a solution $(x(\tau), y(\tau))$ that converges to a solution of the nonlinear complementarity problem. The properties of the path $(x(\tau), y(\tau))$ have been explored in a series of papers by Kojima, Mizuno, and Noma [25], Kojima, Megiddo, and Noma [23], and Kojima, Megiddo, and Mizuno [22]. The continuation approach has received considerable attention in the monotone case since it covers linear programming and convex quadratic programming. Indeed, interior-point methods for linear complementarity problems can be viewed as modified Newton methods that follow the continuation path defined by $(x(\tau), y(\tau))$. See, for example, the discussions in Wright [46] and Kojima, Megiddo, and Mizuno [22]. Recent work on nonlinear monotone mappings includes Chen and Harker [4], Kojima, Noma, and Yoshise [26], and Wright [45].

Disadvantages of the continuation approach are that the method breaks down if f does not satisfy global assumptions that guarantee, in particular, that the solution $(x(\tau), y(\tau))$ of the system (1.4) exists for all $\tau > 0$ and that a starting point x_0 with $x_0 > 0$ and $f(x_0) > 0$ is required. The constrained nonsmooth equations approach, on the other hand, is defined for all mappings f, but is guaranteed to converge to a solution of (1.1) only if the iterates have a limit point that satisfies certain regularity assumptions. A disadvantage of the nonsmooth equations approach is that the lack of differentiability invalidates classical Newton and quasi-Newton methods, and thus it is necessary to device special purpose methods for solving these systems.

Other approaches to the solution of (1.1) include transformation into a smooth system of nonlinear equations (Mangasarian [28], Watson [44], Subramanian [43], Kanzow [21]), and into an optimization problem (Fukushima [13], Mangasarian and Solodov [29], Kanzow [20], Geiger and Kanzow [17]). For additional approaches, see the paper of Harker and Pang [18]. Some of these approaches are related to the constrained nonsmooth equations approach. For example, Fukushima [13] shows that x^* solves the nonlinear complementarity problem (1.1) if and only if x^* is a global solution of

$$\min \left\{ \|f(x)\|_2^2 - \|x - f(x) - [x - f(x)]_+\|_2^2 : x \geq 0 \right\}.$$

Fukushima also shows that a stationary point for this problem is a solution of (1.1) if the Jacobian matrix is positive definite at the stationary point. Global convergence results for these approaches are generally weaker than those for the nonsmooth equations approach, and thus we do not emphasize these results.

We formulate the nonlinear complementarity problem (1.1) as the bound-constrained nonlinear least squares problem

$$\min \left\{ \tfrac{1}{2}\|h(x,y)\|_2^2 : x \geq 0,\ y \geq 0 \right\}. \tag{1.6}$$

We show that this approach is valid because, under reasonable conditions, stationary points of (1.6) are solutions to the nonlinear complementarity problem (1.1). Moreover, we show that our approach is valid under weaker conditions than the nonsmooth equations approach.

We also propose the use of a trust region method for the solution of (1.1). The trust region method is defined for all differentiable mappings f. Moreover, global and superlinear convergence to a solution of (1.1) takes place under reasonable conditions. We also point out that implementation of the trust region method only requires the solution of systems of linear equations; there is no need to solve either a linear programming problem or a quadratic programming problem.

We begin the study of the system of equations defined by (1.4) and (1.5) by introducing the P_0- and the P-functions. These two classes of function were introduced by Moré and Rheinboldt [35] as nonlinear generalizations of the P- and P_0-matrices of Fiedler and Pták [11], and were first used for the study of nonlinear complementarity problems by Moré [33, 34]. The material in Section 2 relates our work to that of Kojima, Megiddo, and Noma [24, 25] and Kojima, Mizuno, and Noma [23] since their convergence results require that f be a P_0-function in $\mathbb{R}^n_+$. As we shall see, the convergence results for the nonsmooth systems approach, and the approach based on (1.6) are also related to the P_0- and the P-functions.

In Section 3 we introduce a regularity assumption, and show that if x^* is regular, then any stationary point (x^*, y^*) of (1.6) produces a solution x^* of the nonlinear complementarity problem (1.1). Our objective is to provide conditions that guarantee that any limit point of a sequence generated by an algorithm for (1.6) is a solution of (1.1). Conditions that guarantee regularity are explored in Section 4. Our results are similar to those obtained by Pang and Gabriel [39], Gabriel and Pang [14, 15], Monteiro, Pang, and Wang [32], and Xiao and Harker [47, 48], but we do not require any explicit assumptions of nonsingularity.

The global convergence of the trust region method is presented in Section 5, with the rate of convergence analysis in Section 6. We show that under reasonable regularity assumptions, the trust region method converges to a solution of the nonlinear complementarity problem (1.1). The rate of convergence is Q-linear or Q-superlinear, depending on the choice of tolerances for the approximate solution of the subproblems. This analysis is of interest because the nonlinear least squares problem (1.6) is degenerate from an optimization viewpoint.

In this paper $\|\cdot\|$ is the Euclidean norm. Vector inequalities apply to each component, and thus $x \geq 0$ if all the components of x are nonnegative. The set $\mathbb{R}^n_+$ consists of all $x \in \mathbb{R}^n$ with $x \geq 0$, and $\mathbb{R}^n_{++}$ is the set of all $x > 0$. Given an index set $\mathcal{C}$, the vector $x_\mathcal{C}$ consists of all components x_i of x with $i \in \mathcal{C}$. For a matrix $A \in \mathbb{R}^{n \times n}$, we use the notation $A_\mathcal{C}$ for the principal submatrix of A with elements $a_{i,j}$ and $i, j \in \mathcal{C}$.

2 Global properties

The P- and P_0-matrices of Fiedler and Pták [11] play a central role in the study of the linear complementarity problem since $A \in \mathbb{R}^n$ is a P-matrix if and only if the linear complementarity problem

$$x \geq 0, \qquad Ax + q \geq 0, \qquad x^T(Ax + q) = 0$$

has a unique solution for any $q \in \mathbb{R}^n$. The nonlinear generalization of the P- and P_0-matrices are the P- and P_0-functions. In this section we show, in particular, that $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a P_0-function on $\mathbb{R}^n_+$ if and only if the function $h : \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ defined by (1.5) is a P_0-function on $\mathbb{R}^{2n}_+$.

A function $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a P-function in a set Ω if for each x and y in Ω with $x \neq y$, there is an index i such that

$$(x_i - y_i)[f_i(x) - f_i(y)] > 0.$$

Similarly, $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a P_0-function in Ω if for each x and y in Ω with $x \neq y$, there is an index i such that

$$x_i \neq y_i, \qquad (x_i - y_i)[f_i(x) - f_i(y)] \geq 0.$$

The P and P_0-functions share many of the properties of P and P_0-matrices. For example, every P-function is also a P_0-function. Moreover, f is a P_0-function if and only if the function $f_\epsilon : \mathbb{R}^n \mapsto \mathbb{R}^n$ defined by $f_\epsilon(x) = f(x) + \epsilon x$ is a P-function for all $\epsilon > 0$. For additional properties of these functions, see Moré and Rheinboldt [35].

Theorem 2.1 *The function $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a P_0-function on $\mathbb{R}^n_+$ if and only if the function $h : \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ is a P_0-function on $\mathbb{R}^{2n}_+$.*

Proof. If h is a P_0-function, then given $\bar{x} \neq x$, there is an index i such that $\bar{x}_i \neq x_i$ and

$$(\bar{x}_i - x_i)[h_i(\bar{x}, 0) - h_i(x, 0)] \geq 0.$$

Since $f_i(x) = h_i(x, 0)$ for $1 \leq i \leq n$, we obtain that

$$(\bar{x}_i - x_i)[f_i(\bar{x}) - f_i(x)] = (\bar{x}_i - x_i)[h_i(\bar{x}, 0) - h_i(x, 0)] \geq 0,$$

and thus f is a P_0-function.

Assume now that f is a P_0-function, and consider two pairs $(\bar{x}, \bar{y}) \neq (x, y)$. First consider the case where $\bar{x} \neq x$. In this case the definition of a P_0-function implies that there is an index i with $\bar{x}_i \neq x_i$ and

$$(\bar{x}_i - x_i)[f_i(\bar{x}) - f_i(x)] \geq 0.$$

If $(\bar{x}_i - x_i)(\bar{y}_i - y_i) \leq 0$, then

$$(\bar{x}_i - x_i)[h_i(\bar{x}, \bar{y}) - h_i(x, y)] = (\bar{x}_i - x_i)[f_i(\bar{x}) - f_i(x)] - (\bar{x}_i - x_i)(\bar{y}_i - x_i) \geq 0.$$

If, on the other hand, $(\bar{x}_i - x_i)(\bar{y}_i - y_i) > 0$, then

$$(\bar{y}_i - y_i)[\bar{y}_i\bar{x}_i - y_i x_i] = (\bar{y}_i - y_i)[(\bar{y}_i - y_i)\bar{x}_i + (\bar{x}_i - x_i)y_i] \geq 0.$$

This yields the desired result if $\bar{x} \neq x$. If $\bar{x} = x$, then

$$(\bar{y}_i - y_i)[\bar{y}_i\bar{x}_i - y_i x_i] = (\bar{y}_i - y_i)^2 x_i \geq 0$$

for all indices. This proves that h is a P_0-function. ■

For differentiable mappings, Moré and Rheinboldt [35] proved that f is a P_0-function in a rectangle Ω if and only if $f'(x)$ is a P_0-matrix for each $x \in \Omega$. We can combine this result with Theorem 2.1 to obtain that if f is differentiable on $\mathbb{R}^n_+$, then h is a P_0-function on $\mathbb{R}^{2n}_+$ if and only if $f'(x)$ is a P_0-matrix for each $x \in \mathbb{R}^n_+$. Stronger results can be obtained if we restrict attention to $\mathbb{R}^n_{++}$.

Theorem 2.2 *If $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is differentiable at $x \in \mathbb{R}^n_+$ and $f'(x)$ is a P_0-matrix, then $h'(x, y)$ is a P_0-matrix for any $y \in \mathbb{R}^n_+$. Moreover, if $y > 0$, then $h'(x, y)$ is nonsingular.*

Proof. The proof is along the same lines as that of Theorem 2.1. We need to show that for any nonzero vector $w \in \mathbb{R}^{2n}$, there is an index i such that

$$u_i \neq 0, \quad u_i[h'(x,y)w]_i \geq 0, \quad \text{or} \quad v_i \neq 0, \quad v_i[h'(x,y)w]_{i+n} \geq 0,$$

where $w = (u, v)$. Assume that $u \neq 0$. Since $f'(x)$ is a P_0-matrix there is an index i with $u_i \neq 0$ such that $u_i[f'(x)u]_i \geq 0$. If $u_i v_i \leq 0$, then

$$u_i[h'(x,y)w]_i = u_i[f'(x)u]_i - u_i v_i \geq 0.$$

If, on the other hand, $u_i v_i > 0$, then

$$v_i[h'(x,y)w]_{i+n} = v_i[y_i u_i + x_i v_i] = y_i u_i v_i + x_i v_i^2 \geq 0.$$

This yields the desired result if $u \neq 0$. If $u = 0$, then

$$v_i[h'(x,y)w]_{i+n} = x_i v_i^2 \geq 0$$

for all indices. This proves that $h'(x,y)$ is a P_0-matrix.

Assume now that $y > 0$ and that $h'(x,y)w = 0$. This implies that

$$f'(x)u = v, \qquad y_i u_i + x_i v_i = 0.$$

If $u \neq 0$, then there is an index i with $u_i \neq 0$ such that $u_i[f'(x)u]_i \geq 0$. Hence, $u_i v_i \geq 0$. This is not possible because $y_i u_i + x_i v_i = 0$, $y_i > 0$, and $u_i \neq 0$. Hence, we must have $u = 0$, and thus $v = f'(x)u = 0$. This proves that $h'(x,y)$ is nonsingular. ■

Theorem 2.2 is closely related to Lemma 5.4 in the paper of Kojima, Megiddo, and Noma [23]. However, they assume that $x > 0$ and do not show that $h'(x,y)$ is a P_0-matrix.

In our situation Theorem 2.2 shows that if f is a P_0-function on $\mathbb{R}^n_+$, then (1.6) has no statioanry points in the interior of $\mathbb{R}^{2n}_+$. Indeed, if (x,y) is a stationary point with $x > 0$ and $y > 0$, then

$$h'(x,y)^T h(x,y) = 0,$$

and since Theorem 2.2 shows that $h'(x,y)$ is nonsingular, we must have $h(x,y) = 0$. In particular, $x_i y_i = 0$. This contradicts our assumption that (x,y) is in the interior of $\mathbb{R}^{2n}_+$.

Theorem 2.3 *The function $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a P-function on $\mathbb{R}^n_{++}$ if and only if the function $h : \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ is a P-function on $\mathbb{R}^{2n}_{++}$.*

Proof. The proof is very similar to that of Theorem 2.1. ■

Moré and Rheinboldt [35] proved that if f is differentiable on a rectangle Ω and $f'(x)$ is a P-matrix for each $x \in \Omega$, then f is a P-function in Ω; the converse is clearly false. Combining this result with Theorem 2.3 we obtain that if f is differentiable on $\mathbb{R}^n_+$ and $f'(x)$ is a P-matrix for each $x \in \mathbb{R}^{2n}_+$, then h is a P-function on $\mathbb{R}^{2n}_{++}$.

3 Regularity

We want to show that if x^* satisfies a regularity assumption, then any stationary point (x^*, y^*) of (1.6) yields a solution x^* to the nonlinear complementarity problem (1.1). The regularity assumption requires the introduction of index sets associated with a vector x in $\mathbb{R}^n_+$. The first index set

$$\mathcal{C} = \{i : x_i \geq 0,\ f_i(x) \geq 0,\ x_i f_i(x) = 0\}$$

is the set of indices that are complementary ($\mathcal{C}$ for complementary). Other indices can be classified according to the sign of $x_i f_i(x)$; we have the negative and positive indices

$$\mathcal{N} = \{i : x_i > 0,\ f_i(x) < 0\}, \qquad \mathcal{P} = \{i : x_i > 0,\ f_i(x) > 0\},$$

and the residual indices

$$\mathcal{R} = \{i : x_i = 0,\ f_i(x) < 0\}.$$

Note that these sets depend on a given $x \in \mathbb{R}^n_+$, but that the notation does not reflect this dependence. This should not cause any confusion because the given $x \in \mathbb{R}^n_+$ will always be clear from the context.

Definition 3.1 *The vector $x \in \mathbb{R}^n_+$ is regular with respect to problem* (1.6) *if for any nonzero $z \in \mathbb{R}^n$ such that*

$$z_{\mathcal{C}} = 0, \qquad z_{\mathcal{N}} < 0, \qquad z_{\mathcal{P}} > 0, \qquad z_{\mathcal{R}} < 0, \tag{3.1}$$

there is a vector $p \in \mathbb{R}^n$ such that

$$p_{\mathcal{C}} \leq 0, \qquad p_{\mathcal{P}} \geq 0, \qquad p_{\mathcal{R}} \leq 0, \tag{3.2}$$

with

$$z^T f'(x) p > 0. \tag{3.3}$$

Our objective is to provide conditions that guarantee that any limit point of a sequence generated by an algorithm for (1.6) is a solution of (1.1). Thus we are only interested in the regularity of points of attraction for a minimization algorithm for (1.6)

A solution of the nonlinear complementarity problem (1.1) is regular according to this definition. This is clear because at a solution of (1.1) the sets $\mathcal{P}$, $\mathcal{N}$, and $\mathcal{R}$ are empty, and thus the only vector z that satisfies (3.1) is the zero vector.

Also note that the regularity condition imposes a restriction on the rows of the Jacobian matrix $f'(x)$ with indices in

$$\mathcal{D} = \mathcal{N} \cup \mathcal{P} \cup \mathcal{R}.$$

The set $\mathcal{D}$ is the set of defective indices; indices in the set $\mathcal{C}$ are not relevant because $z_{\mathcal{C}} = 0$. We usually choose $p_{\mathcal{C}} = 0$ and guarantee regularity by imposing conditions on $[f'(x)]_{\mathcal{D}}$. For example, note that if we assume that $[f'(x)]_{\mathcal{D}}$ is positive definite, then x is regular because we can choose $p = z$. In particular, x is regular if $f'(x)$ is positive definite.

Theorem 3.2 *Assume that $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is differentiable on $\mathbb{R}^n_+$. If (x^*, y^*) is a stationary point of* (1.6)*, then x^* is regular if and only if x^* solves the nonlinear complementarity problem* (1.1)*.*

Proof. We have already noted that if x^* is a solution to the nonlinear complementarity problem (1.1), then x^* is regular, so we only need to prove the converse.

If (x, y) is a stationary point of the nonlinear least squares problem (3.2), then the Kuhn-Tucker conditions imply that there are multipliers u and v such that

$$\begin{pmatrix} f'(x)^T & Y \\ -I & X \end{pmatrix} \begin{pmatrix} f(x) - y \\ Yx \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix},$$

where u and v satisfy the complementarity conditions

$$u \geq 0, \qquad v \geq 0, \qquad u^T x = 0, \qquad v^T y = 0.$$

We express these conditions in terms of vectors $z = f(x) - y$, and $w = Y^2 x$ as

$$f'(x)^T z + w = u, \qquad X^2 y - z = v. \tag{3.4}$$

The first relation will not be needed until the end of the proof. However, in the proof below we make heavy use of the relationships $x_i^2 y_i - z_i = v_i$ and $z_i = f_i(x) - y_i$.

First consider the case where $z = 0$. If $z = 0$, then $f(x) = y \geq 0$. Moreover, $x_i^2 y_i = v_i$, and since $v_i y_i = 0$, we must have $x_i y_i = 0$. Hence, x solves (1.1). In the remainder of the proof we show that the regularity assumption implies that $z = 0$.

The first step in the proof is to show that

$$z_{\mathcal{C}} = 0, \qquad z_{\mathcal{N}} < 0, \qquad z_{\mathcal{P}} > 0, \qquad z_{\mathcal{R}} < 0. \tag{3.5}$$

We first show that $z_{\mathcal{C}} = 0$. If $i \in \mathcal{C}$, then $x_i^2 y_i - z_i = v_i$ shows that if $x_i = 0$, then $z_i = -v_i \leq 0$, while $z_i = f_i(x) - y_i$ shows that if $f_i(x) = 0$, then $z_i = -y_i \leq 0$. Hence, $z_i \leq 0$. If we assume that $z_i < 0$, then $v_i = x_i^2 y_i - z_i > 0$, and thus $y_i = 0$ by the complementarity conditions. This leads to the contradiction $z_i = f_i(x) - y_i \geq 0$. Hence, $z_i = 0$ in this case.

We prove that $z_{\mathcal{P}} > 0$ by considering two cases. If $i \in \mathcal{P}$ and $v_i = 0$, then $z_i = x_i^2 y_i$. We cannot have $y_i = 0$ because then $z_i = 0$, but on the other hand, $z_i = f_i(x) > 0$ for $i \in \mathcal{P}$. This contradiction shows that $y_i > 0$, and thus $z_i > 0$. If $i \in \mathcal{P}$ and $v_i > 0$, then $y_i = 0$ by complementarity, and thus $z_i = f_i(x) > 0$. Hence, we have shown that $z_{\mathcal{P}} > 0$.

The proof that $z_{\mathcal{N}} < 0$ and that $z_{\mathcal{R}} < 0$ follows directly by noting that $f_i(x) < 0$ for $i \in \mathcal{N} \cup \mathcal{R}$, that $y_i \geq 0$, and that $z_i = f_i(x) - y_i$.

The next step in the proof is to show that

$$w_{\mathcal{C}} = 0, \qquad w_{\mathcal{N}} = 0, \qquad w_{\mathcal{R}} = 0. \tag{3.6}$$

We first prove that $w_{\mathcal{C}} = 0$. Assume that $i \in \mathcal{C}$. If $x_i = 0$, then $w_i = 0$ is immediate. If $x_i > 0$, then $f_i(x) = 0$, and since we have already shown that $z_i = 0$, we obtain that $y_i = 0$. Hence, $w_i = 0$ as desired.

We prove that $w_{\mathcal{N}} = 0$ by first recalling that $x_i^2 y_i = z_i + v_i$ and that $z_i < 0$ for $i \in \mathcal{N}$. Hence, $x_i^2 y_i < v_i$, and since $y_i v_i = 0$, we must have $y_i = 0$. Hence, $w_{\mathcal{N}} = 0$. The proof that $w_{\mathcal{R}} = 0$ is immediate since $x_i = 0$ for $i \in \mathcal{R}$.

We also need to note that since $x_i > 0$ for $i \in \mathcal{N} \cup \mathcal{P}$, the complementarity conditions show that

$$u_{\mathcal{N}} = 0, \qquad u_{\mathcal{P}} = 0. \tag{3.7}$$

For the final step in the proof, let p be the vector guaranteed by the regularity assumption, and assume that $z \neq 0$. The results (3.6) and (3.7), together with (3.2), clearly show that

$$w^T p \geq 0, \qquad u^T p \leq 0.$$

The Kuhn-Tucker conditions (3.4) imply that $f'(x)^T z + w = u$, and thus these inequalities show that $z^T f'(x) p \leq 0$. This contradicts (3.3) and proves our result. ■

Pang and Gabriel [39] proved a result similar to Theorem 3.2. They considered the problem

$$\min \left\{ \tfrac{1}{2} \|H(x)\|^2 : x \geq 0 \right\}, \tag{3.8}$$

where

$$H(x) = \min(x, f(x)),$$

and showed that $x \in \mathbb{R}^n_+$ is a stationary point for this problem if and only if $x \in \mathbb{R}^n_+$ is s-regular, in the sense that there is a $p \in \mathbb{R}^n$ such that

$$\begin{array}{lll} [f'(x)p]_i = -f_i(x) & \text{if} & f_i(x) < x_i, \ x_i > 0 \\ [f'(x)p]_i \geq -f_i(x) & \text{if} & f_i(x) < x_i, \ x_i = 0 \\ [f'(x)p]_i \leq -f_i(x) & \text{if} & f_i(x) = x_i, \ x_i > 0 \end{array}$$

and

$$\begin{array}{lll} p_i \geq -x_i & \text{if} & f_i(x) < x_i, \ x_i = 0 \\ p_i \leq -x_i & \text{if} & f_i(x) = x_i, \ x_i > 0 \\ p_i = -x_i & \text{if} & f_i(x) > x_i \text{ or } f_i(x) = x_i, \ x_i = 0. \end{array}$$

A difference between our regularity assumption and s-regularity is that s-regularity requires that the vector p satisfy the equations

$$[f'(x)p]_i = -f_i(x), \qquad i \in \mathcal{I} = \{i : f_i(x) < x_i, \ x_i > 0\} .$$

This can be satisfied for arbitrary f only if we assume that $[f'(x)]_{\mathcal{I}}$ is nonsingular.

Also note that the notion of s-regularity depends on the scaling of f; that is, if we consider the scaled function $f_s(x) = \alpha f(x)$ where $\alpha > 0$, then f may be s-regular at a given x, but f_s can fail to be s-regular at x. The dependence of s-regularity on the scaling can be explained by noting that problem (3.8) depends on the scaling of f. On the other hand, a computation shows that the stationary points of (1.6) are unchanged under this change of scale.

We should also consider more general scalings since the nonlinear complementarity problem (1.1) is invariant under the change of scale defined by $f_s(x) = D_r f(D_v x)$, where D_r and D_v are diagonal matrices with positive diagonal entries. Note, however, that stationary points of (1.6) are not invariant under this change of scale unless we replace the Euclidean norm by a scaled norm that reflects the scaling in the problem.

An advantage of our approach is that most algorithms for solving the minimization problem (1.6) can be shown to generate sequences $\{(x_k, y_k)\}$ such that any limit point (x^*, y^*) of $\{(x_k, y_k)\}$ is a stationary point of (1.6). Hence, Theorem 3.2 shows that if x^* is regular, then x^* is a solution of the nonlinear complementarity problem (1.1). The assumptions needed to obtain this result are usually that f is continuously differentiable on $\mathbb{R}^n_+$ and that the level set

$$\{(x, y) \in \mathbb{R}^n_+ : \|h(x, y)\| \le \|h(x_0, y_0)\|\}$$

is bounded. Pang and Gabriel [39], and Gabriel and Pang [14] can show that limit points are solutions of the nonlinear complementarity problem (1.1) only if they assume that x^* is s-regular and $[f'(x^*)]_{\mathcal{I}}$ is nonsingular for any index set $\mathcal{I}$ such that

$$\{i : f_i(x^*) < x_i^*, \ x_i^* > 0\} \subset \mathcal{I} \subset \{i : f_i(x^*) \le x_i^*\} .$$

In their terminology, they need to assume that the limit x^* is b-regular.

The regularity assumption needed by the algorithm of Xiao and Harker [47, 48] is similar in the sense that limit points of the sequence generated by their algorithm are guaranteed to be solutions of (1.1) only if $[f'(x^*)]_{\mathcal{I}}$ is nonsingular for any index set $\mathcal{I}$ such that

$$\{i : x_i^* > 0\} \subset \mathcal{I} \subset \{i : x_i^* \ge 0\} .$$

We provide a precise definition of this regularity assumption at the end of the next section.

4 Regularity Conditions

In this section we explore conditions that guarantee regularity of x in terms of the Jacobian matrix $f'(x)$. This requires the introduction of two classes of matrices that have played an important role in the study of complementarity problems.

A matrix $A \in \mathbb{R}^{n\times n}$ is a P-matrix if for each $x \neq 0$ there is an index i such that $x_i[Ax)]_i > 0$. Similarly, a matrix $A \in \mathbb{R}^{n\times n}$ is a P_0-matrix if for each $x \neq 0$ there is an index i such that $x_i \neq 0$ and $x_i[Ax]_i \geq 0$.

The P-matrices were introduced by Fiedler and Pták [11] as generalizations of the positive definite matrices, the M-matrices, and the irreducibly diagonally dominant matrices. One of the main reasons for the importance of the P-matrices in the study of linear complementarity problems is that $A \in \mathbb{R}^n$ is a P-matrix if and only if the linear complementarity problem

$$x \geq 0, \qquad Ax + q \geq 0, \qquad x^T(Ax+q) = 0$$

has a unique solution for any $q \in \mathbb{R}^n$.

Fiedler and Pták [12] also defined the S-matrices: A matrix $A \in \mathbb{R}^{n\times n}$ is an S-matrix if there is an $x \neq 0$ such that $x \geq 0$ and $Ax > 0$, while $A \in \mathbb{R}^{n\times n}$ is an S_0-matrix if there is an $x \neq 0$ such that $x \geq 0$ and $Ax \geq 0$. Clearly, A is an S-matrix if and only if the above linear complementarity is feasible for all $q \in \mathbb{R}^n$.

The P-matrices and S-matrices are related. Indeed, a P-matrix must be an S-matrix, and any P_0-matrix must be an S_0-matrix. This result of Fiedler and Pták [12] is a direct consequence of the following classical theorem of the alternative.

Theorem 4.1 *Let $A \in \mathbb{R}^{m\times n}$.*

A is an S-matrix if and only if $\left\{ y : y \geq 0,\ A^T y \leq 0,\ y \neq 0 \right\}$ is empty.

A is an S_0-matrix if and only if $\left\{ y : y \geq 0,\ A^T y < 0 \right\}$ is empty.

For additional information on P- and S-matrices and their connection to linear complementarity problems, see the book of Cottle, Pang, and Stone [8].

The regularity requirements can be expressed in terms of nonnegative vectors by making a transformation. Let $z \in \mathbb{R}^n$ be a nonzero vector that satisfies (3.1), and let T be the nonsingular diagonal matrix defined by

$$T = \operatorname{diag}(t_i), \qquad t_i = \begin{cases} 1 & i \in \mathcal{P} \\ -1 & i \notin \mathcal{P} \end{cases} \tag{4.1}$$

Define the transformed z by $\tilde{z} = Tz$, and note that $\tilde{z}_{\mathcal{C}} = 0$ and that $\tilde{z}_{\mathcal{D}} > 0$. We carry out the same transformation on $f'(x)$ and define a transformed Jacobian matrix

$$J(x) = T^{-1} f'(x)\, T^{-1}. \tag{4.2}$$

With this transformation, x is regular if there is a vector $p \in \mathbb{R}^n$ such that $\tilde{p} = Tp$ satisfies

$$\tilde{z}^T J(x)\tilde{p} > 0, \qquad \tilde{p}_i \geq 0, \ i \notin \mathcal{N}. \tag{4.3}$$

The following result is a direct consequence of this observation.

Theorem 4.2 *Assume that $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is differentiable at $x \in \mathbb{R}^n_+$ and let the matrix $J(x)$ be defined by* (4.1) *and* (4.2). *If $[J(x)]_{\mathcal{E}}$ is an S-matrix for some index set $\mathcal{E}$ with $\mathcal{D} \subset \mathcal{E}$, then x is regular.*

Proof. If $[J(x)]_{\mathcal{E}}$ is an S-matrix, then a computation shows that there is a vector $\tilde{p} \in \mathbb{R}^n$ such that $\tilde{p} > 0$ and $[J(x)\tilde{p}]_{\mathcal{E}} > 0$. Since $\tilde{z}_{\mathcal{C}} = 0$, $\tilde{z}_{\mathcal{D}} > 0$, and $\mathcal{D} \subset \mathcal{E}$, it is clear that $\tilde{p}$ satisfies (4.3). Hence, x is regular. ∎

Theorem 4.2 is easy to apply is specific cases. For example, if $x > 0$ and $f(x) > 0$, then $J(x) = f'(x)$, and thus Theorem 4.2 shows that x is regular if $f'(x)$ is an S-matrix. We illustrate this remark with a nonlinear complementarity problem proposed by Kojima (see, for example, Dirkse and Ferris [10]) where f is defined by

$$f(x) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 3x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 3x_4 - 1 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix}. \tag{4.4}$$

Xiao and Harker [47] noted that their damped Newton method converges to x^*, with

$$x^* = (1.05, 1.33, 1.26, 0.0), \qquad f(x^*) = (4.98, 6.86, 9.84, 5.99),$$

if started from $x_0 = (1, 1, 1, 1)$. The damped Newton method fails at x^* because the (linear complementarity) subproblem required by this method is infeasible. On the other hand, x^* cannot be a limit point for the trust region Newton method that we propose in Section 5 (or any reasonable algorithm based on the formulation (1.6)) because x^* is regular. A full justification of this remark is provided by the results in Section 5, but for now note that $f'(x^*)$ is an S-matrix, and thus Theorem 4.2 shows that x^* is regular.

Theorem 4.2 is also easy to apply if $x > 0$ and $f(x) < 0$. In this case $J(x) = f'(x)$, and thus Theorem 4.2 shows that x is regular if $f'(x)$ is an S-matrix. This remark applies to the problem defined by (4.4), because for this problem $f'(x)$ is an S-matrix.

Two special cases of Theorem 4.2 are of interest. For the following result note that the definition of a P-matrix implies that if D is any nonsingular diagonal matrix, then A is a P-matrix (P_0-matrix) if and only if DAD is a P-matrix (P_0-matrix).

Corollary 4.3 *If $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is differentiable at $x \in \mathbb{R}^n_+$ and $[f'(x)]_{\mathcal{D}}$ is a P-matrix, then x is regular.*

Proof. The result follows from Theorem 4.2 because if $[f'(x)]_{\mathcal{D}}$ is a P-matrix, then $[J(x)]_{\mathcal{D}}$ is also a P-matrix, and thus an S-matrix. ■

Corollary 4.3 shows, in particular, that if $f'(x)$ is a positive definite matrix, an M-matrix, or an H-matrix with positive diagonal entries ($f'(x)$ is an H-matrix if $f'(x)D$ is strictly diagonally dominant for some diagonal matrix D with positive diagonal entries), then x is regular.

We now consider the positive semidefinite case in more detail. As mentioned in the introduction, this case is of special interest because it covers the linear programming and convex quadratic programming problems.

Corollary 4.4 *Assume that $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is differentiable at $x \in \mathbb{R}^n_+$ and that (x, y) is a stationary point of* (1.6) *for some $y \in \mathbb{R}^n_+$. If $[f'(x)]_{\mathcal{D}}$ is positive semidefinite, then $J(x) = f'(x)$. If, in addition, $[f'(x)]_{\mathcal{E}}$ is an S-matrix for some index set $\mathcal{E}$ with $\mathcal{D} \subset \mathcal{E}$, then x is regular.*

Proof. We first show that if $[f'(x)]_{\mathcal{D}}$ is positive semidefinite, then $\mathcal{P}$ is empty. The proof uses results and notation established during the proof of Theorem 3.2. In particular, recall that we have already shown that if $\mathcal{P}$ is not empty, then $z_{\mathcal{P}} > 0$.

We prove that $w_{\mathcal{P}} > 0$ by noting that $w_i = y_i^2 x_i$ and $x_i > 0$ for $i \in \mathcal{P}$ and that $y_i = z_i + f_i(x) > 0$ for $i \in \mathcal{P}$. Now recall (3.5), (3.6), (3.7), and that the Kuhn-Tucker conditions (3.4) imply that $f'(x)^T z + w = u$ with $u \geq 0$. Hence,

$$z^T f'(x) z = z_{\mathcal{D}}^T f'(x) z_{\mathcal{D}} = u_{\mathcal{R}}^T z_{\mathcal{R}} - w_{\mathcal{P}}^T z_{\mathcal{P}} \leq -w_{\mathcal{P}}^T z_{\mathcal{P}} < 0.$$

This contradicts the assumption that $[f'(x)]_{\mathcal{D}}$ is positive semidefinite. Hence, $\mathcal{P}$ is empty, and thus $J(x) = f'(x)$.

If we also assume that $[f'(x)]_{\mathcal{E}}$ is an S-matrix, then $[J(x)]_{\mathcal{E}}$ is an S-matrix, and the result follows from Theorem 4.2. ■

In this result we assumed that (x, y) is a stationary point of (1.6) for some $y \in \mathbb{R}^n_+$. There is no loss of generality in assuming this because we are only interested in the regularity of points of attraction for an algorithm for (1.6).

Corollary 4.4 shows that we can guarantee the regularity of x without imposing any nonsingularity assumptions on $f'(x)$. For example, if $f'(x) = u(x)u(x)^T$ for any $u(x) > 0$, then $f'(x)$ is a positive semidefinite S-matrix. On the other hand, $f'(x)$ is clearly singular.

Theorem 4.5 *If $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is differentiable at $x \in \mathbb{R}^n_+$ and $[f'(x)]_{\mathcal{D}}$ is a nonsingular P_0-matrix, then x is regular.*

Proof. Since $[f'(x)]_{\mathcal{D}}$ is a P_0-matrix, $[J(x)]_{\mathcal{D}}$ is also a P_0-matrix, and thus an S_0-matrix. Hence, there is a vector $\tilde{p} \neq 0$, with $\tilde{p}_{\mathcal{C}} = 0$ and $\tilde{p}_{\mathcal{D}} \geq 0$ such that $[J(x)]_{\mathcal{D}} \tilde{p}_{\mathcal{D}} \geq 0$. Since

$\tilde{z}_{\mathcal{C}} = 0$ and $\tilde{z}_{\mathcal{D}} > 0$, we can fail to satisfy (4.3) only if $[J(x)]_{\mathcal{D}} \tilde{p}_{\mathcal{D}} = 0$. However, this is not possible because $[J(x)]_{\mathcal{D}}$ is nonsingular and $\tilde{p}_{\mathcal{D}} \neq 0$. Hence, x is regular. ■

The main assumption used in the continuation approach of Kojima, Megiddo, and Noma [23] is that f is a P_0-function. Under this assumption Moré and Rheinboldt [35] proved that $f'(x)$ is a P_0-matrix. Thus, Theorem 4.5 shows that if we restrict ourselves to P_0-functions, then we need to assume that $[f'(x)]_{\mathcal{D}}$ is nonsingular to guarantee that an algorithm based on the formulation (1.6) converges to a solution of (1.1).

We now present a variation on the previous results that is closely related to the results of Pang and Gabriel [39], Gabriel and Pang [14], and Xiao and Harker [47, 48]. Recall that if $A \in \mathbb{R}^{n \times n}$ is partitioned in the form

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} \tag{4.5}$$

and the matrix $A_{1,1}$ is nonsingular, then $A_S = A_{2,2} - A_{2,1} A_{1,1}^{-1} A_{1,2}$ is the Schur complement of $A_{1,1}$ in A.

Theorem 4.6 *Let $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ be differentiable at $x \in \mathbb{R}^n_+$. If $[f'(x)]_{\mathcal{N}}$ is nonsingular and the Schur complement of $[f'(x)]_{\mathcal{N}}$ in $[J(x)]_{\mathcal{D}}$ is an S-matrix, then x is regular.*

Proof. Partition $[J(x)]_{\mathcal{D}}$ into

$$[J(x)]_{\mathcal{D}} = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix},$$

where $A_{1,1} = [f'(x)]_{\mathcal{N}}$. We want to show that there are vectors p_1 and p_2 such that

$$\begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 0 \\ q \end{pmatrix}, \qquad p_2 > 0, \qquad q > 0.$$

Since $A_{1,1}$ is nonsingular, this system is equivalent to

$$\begin{pmatrix} A_{1,1} & A_{1,2} \\ 0 & A_S \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 0 \\ q \end{pmatrix},$$

where $A_S = A_{2,2} - A_{2,1} A_{1,1}^{-1} A_{1,2}$ is the Schur complement of $[f'(x)]_{\mathcal{N}}$ in $[J(x)]_{\mathcal{D}}$. Since A_S is an S-matrix, we can find a vector $p_2 > 0$ such that $A_S p_2 > 0$, and since $A_{1,1} = [f'(x)]_{\mathcal{N}}$ is nonsingular, we can solve $A_{1,1} p_1 + A_{1,2} p_2 = 0$ for p_1. Hence,

$$[J(x)]_{\mathcal{D}} \tilde{p}_{\mathcal{D}} = \begin{pmatrix} 0 \\ q \end{pmatrix}, \qquad \tilde{p}_{\mathcal{D}} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix},$$

satisfies (4.3) because $\tilde{z}_{\mathcal{C}} = 0$ and $\bar{z}_{\mathcal{D}} > 0$. Hence, x is regular. ■

Corollary 4.6 is similar to a result of Pang and Gabriel [39] that guarantees s-regularity. In their Proposition 3, they assume that $[f'(x)]_{\mathcal{N}}$ is nonsingular and that the Schur complement of $[f'(x)]_{\mathcal{N}}$ in $[f'(x)]_{\mathcal{D}}$ is an S-matrix, but the definitions of the index sets $\mathcal{N}$ and $\mathcal{D}$ are different. For example, in their result

$$\mathcal{D} = \{i : f_i(x) < x_i,\ x_i \geq 0\}, \qquad \mathcal{N} = \{i : f_i(x) < x_i,\ x_i > 0\}.$$

For additional details, see Pang and Gabriel [39].

Corollary 4.6 is also related to the definition of regularity used by Ralph [40], Dirkse and Ferris [9], and Xiao and Harker [47, 48]. They require that $[f'(x)]_{\mathcal{K}}$ be nonsingular, where

$$\mathcal{K} = \{i :\ x_i > 0\},$$

and that the Schur complement of $[f'(x)]_{\mathcal{K}}$ in $[f'(x)]_{\mathcal{L}}$ be a P-matrix, where

$$\mathcal{L} = \{i :\ x_i \geq 0\}.$$

These conditions imply that $[f'(x)]_{\mathcal{I}}$ is nonsingular for any index set $\mathcal{I}$ with $\mathcal{K} \subset \mathcal{I} \subset \mathcal{L}$. In particular, $[f'(x)]_{\mathcal{L}}$ must be nonsingular.

We cannot compare their results with ours because the assumptions are made on different submatrices of the Jacobian matrix. The major difference seems to be that results based on our regularity condition do not require explicit nonsingularity assumptions.

5 Global Convergence

We now show that a trust region method for the solution of (1.6), where $h : \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ is defined by (1.5), can be used to generate a sequence $\{(x_k, y_k)\}$ such that if $\{x_k\}$ has a limit point x^* that is regular, then x^* is a solution to the nonlinear complementarity problem (1.1). We use the trust region method of Burke, Moré, and Toraldo [3] for the general minimization problem

$$\min \{f_0(z) : z \in \Omega\}, \tag{5.1}$$

where Ω is a general closed convex set, but specialized to the case where

$$f_0(z) = \tfrac{1}{2}\|h(z)\|^2, \qquad \Omega = \mathbb{R}^{2n}_+. \tag{5.2}$$

Other algorithms could have been used, for example, those of Bertsekas [2], Gay [16], Soares and Júdice [42], and Coleman and Li [5], but the rate of convergence theory of these algorithms does not cover degenerate minimization problems. For problem (5.1) with $\Omega = \mathbb{R}^{2n}_+$, nondegeneracy means that

$$[\nabla f_0(z^*)]_i \neq 0, \qquad 1 \leq i \leq 2n. \tag{5.3}$$

However, in our case f_0 is given by (5.2), and $h(z^*) = 0$ because z^* is a solution of the nonlinear complementarity problem. Hence,

$$\nabla f_0(z^*) = h'(z^*)^T h(z^*) = 0,$$

so that the solution is degenerate. The rate of convergence theory for the algorithm of Conn, Gould, and Toint [6, 7] has been extended to the degenerate case by Lescrenier [27], but these results do not cover the projected searches that we are proposing. We discuss this point further in the next section.

The trust region method generates a sequence $\{z_k\}$, where $z_k = (x_k, y_k)$ and $z_k \in \mathbb{R}^{2n}_+$. At each iteration we have a bound Δ_k and a model $\psi_k : R^n \to R$ of the possible reduction $f_0(z_k + w) - f_0(z_k)$ for $\|w\| \le \Delta_k$. We use

$$\psi_k(w) = \tfrac{1}{2}\left(\|h(z_k) + h'(z_k)w\|^2 - \|h(z_k)\|^2\right).$$

The iterate z_k and the bound Δ_k are updated according to rules that are standard in trust region methods for unconstrained minimization. Given a step s_k such that $z_k + s_k \in \mathbb{R}^{2n}_+$ and $\psi_k(s_k) < 0$, these rules depend on the ratio

$$\rho_k = \frac{f_0(z_k + s_k) - f_0(z_k)}{\psi_k(s_k)} = \frac{\|h(z_k + s_k)\|^2 - \|h(z_k)\|^2}{\|h(z_k) + h'(z_k)s_k\|^2 - \|h(z_k)\|^2} \tag{5.4}$$

of the actual reduction in the function to the predicted reduction in the model. Since the step s_k is chosen so that $\psi_k(s_k) < 0$, a step with $\rho_k > 0$ yields a reduction in the function. Given $\eta_1 > 0$, the iterate x_k is updated by setting

$$z_{k+1} = \begin{cases} z_k + s_k & \text{if} \quad \rho_k > \eta_1 \\ z_k & \text{if} \quad \rho_k \le \eta_1 \end{cases}$$

The updating rules for Δ_k depend on a constant η_2 such that

$$0 < \eta_1 < \eta_2 < 1,$$

while the rate at which Δ_k is either increased or decreased depend on constants σ_1, σ_2, and σ_3 such that

$$0 < \sigma_1 < \sigma_2 < 1 < \sigma_3.$$

The trust region bound Δ_k is updated by setting

$$\begin{array}{lll} \Delta_{k+1} \in [\sigma_1\Delta_k, \sigma_2\Delta_k] & \text{if} & \rho_k \le \eta_1 \\ \Delta_{k+1} \in [\sigma_1\Delta_k, \sigma_3\Delta_k] & \text{if} & \rho_k \in (\eta_1, \eta_2) \\ \Delta_{k+1} \in [\Delta_k, \sigma_3\Delta_k] & \text{if} & \rho_k \ge \eta_2 \end{array}$$

We choose a step s_k that gives as much reduction in the model ψ_k as the Cauchy step s_k^C generated by the gradient projection method applied to the subproblem

$$\min\left\{\psi_k(w) : x_k + w \in \Omega,\ \|w\| \le \Delta_k\right\}.$$

The Cauchy step s_k^C is of the form $s_k(\alpha_k)$, where the function $s_k(\cdot)$ is defined by

$$s_k(\alpha) = [z_k - \alpha \nabla f_0(z_k)]_+ - z_k$$

and the scalar α_k is chosen so that $s_k(\alpha_k)$ produces a sufficient reduction. We require that

$$\psi_k(s_k(\alpha_k)) \le \mu_0 \left(\nabla f_0(z_k)^T s_k(\alpha_k)\right), \qquad \|s_k(\alpha_k)\| \le \mu_1 \Delta_k, \tag{5.5}$$

for constants μ_0 and μ_1 such that

$$0 < \mu_0 < \tfrac{1}{2}, \qquad \mu_1 > 0,$$

and that there are positive constants γ_1 and γ_2 such that

$$\alpha_k \ge \gamma_1 \quad \text{or} \quad \alpha_k \ge \gamma_2 \bar{\alpha}_k,$$

where $\bar{\alpha}_k > 0$ satisfies

$$\psi_k(s_k(\bar{\alpha}_k)) \ge (1 - \mu_0)\left(\nabla f_0(z_k)^T s_k(\bar{\alpha}_k)\right) \quad \text{or} \quad \|s_k(\bar{\alpha}_k)\| \ge \mu_1 \Delta_k.$$

The function defined by $s_k(\cdot)$ defines a piecewise linear path on the feasible set, and the composite function $\psi_k(s_k(\cdot))$ is a piecewise quadratic that is convex on each piece. The above requirements on α_k require that $\psi_k(s_k(\alpha_k))$ achieve a sufficient reduction as compared with the linear model and that the step $s_k(\alpha_k)$ not be too small. For more information on the above requirements on α_k, see Burke, Moré, and Toraldo [3].

These requirements can be satisfied by generating a decreasing sequence $\{\alpha_k^{(l)}\}$ of positive trial values such that

$$\alpha_k^{(l+1)} \in [\beta_1 \alpha_k^{(l)}, \beta_2 \alpha_k^{(l)}], \qquad 0 < \beta_1 < \beta_2 < 1,$$

with $\alpha_k^{(0)}$ bounded away from zero, and setting α_k to the first trial value that satisfies the sufficient decrease condition (5.5). An advantage of this procedure is that it produces an acceptable α_k with a finite number of evaluations of ψ_k. For more details on this type of search, see Section 4 of Moré and Toraldo [36].

Given the Cauchy step s_k^C, we require that the step s_k satisfy

$$\psi_k(s_k) \le \mu_0 \psi_k(s_k^C), \qquad \|s_k\| \le \mu_1 \Delta_k, \qquad z_k + s_k \in \mathbb{R}_+^{2n}. \tag{5.6}$$

This requirement is quite natural and can always be satisfied by choosing $s_k = s_k^C$. However, this choice is likely to lead to slow convergence, since the method would then reduce to a version of steepest descent. In the next section we explore other choices that lead to superlinear and quadratic convergence.

Theorem 5.1 *Let $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ be continuously differentiable on $\mathbb{R}^n_+$, and let $\{(x_k, y_k)\}$ be the sequence generated by the trust region Newton method. Assume that $\{x_k\}$ is bounded, and let x^* be a limit point of $\{x_k\}$. If x^* is regular, then x^* is a solution to the nonlinear complementarity problem* (1.1) *and*

$$\lim_{k \to +\infty} \|h(x_k, y_k)\| = 0.$$

Proof. The trust region method generates a sequence $\{z_k\}$ such that $\{\|h(z_k)\|\}$ is decreasing. Since $\{x_k\}$ is bounded, and

$$\|f(x_k) - y_k\| \le \|h(z_k)\| \le \|h(z_0)\|,$$

the sequence $\{y_k\}$ is bounded. Hence, $\{z_k\}$ is bounded. Theorem 5.4 of Burke, Moré, and Toraldo [3] now implies that every limit point of $\{z_k\}$ is a stationary point of (1.6). Since x^* is regular, Theorem 3.2 shows that x^* solves the nonlinear complementarity problem (1.1). ■

6 Superlinear Convergence

The analysis of the rate of convergence for the trust region method is delicate because, as mentioned in the preceding section, the minimization problem (1.6) is degenerate at a solution of the nonlinear complementarity problem. In this section we show that we can still obtain superlinear convergence if $h'(z^*)$ is nonsingular at a solution of (1.1).

We have already explored in Section 5 conditions that guarantee that a limit point of the sequence generated by the trust region Newton method is a solution of the nonlinear complementarity problem (1.1). In this section we assume that x^* is a solution of (1.1) with $x^* + f(x^*) > 0$ and that

$$[f'(x^*)]_{\mathcal{B}}, \qquad \mathcal{B} = \{i : x_i^* > 0,\ f_i(x^*) = 0\}, \tag{6.1}$$

is nonsingular. These assumptions are reasonable because if we knew $\mathcal{B}$, we would need to solve the system of nonlinear equations

$$f_i(x^*_{\mathcal{B}}, 0) = 0, \qquad i \in \mathcal{B},$$

to determine $x^*_{\mathcal{B}}$. The following result provides additional motivation for our assumptions.

Theorem 6.1 *If $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ be continuously differentiable on $\mathbb{R}^n_+$, and x^* is a solution of the nonlinear complementarity problem* (1.1)*, then $h'(x^*, y^*)$ is nonsingular for $y^* = f(x^*)$ if and only if* (6.1) *is nonsingular and $x^* + f(x^*) > 0$.*

Proof. If $h'(x^*, y^*)$ is nonsingular, then we must have $x_i^* + y_i^* > 0$ for all i; otherwise a row of $h'(x^*, y^*)$ would be zero. Hence, $x^* + f(x^*) > 0$ when $y^* = f(x^*)$. Also note that if u is any vector with $u_i = 0$ for $i \notin \mathcal{B}$, then

$$h'(x^*, y^*) \begin{pmatrix} u \\ 0 \end{pmatrix} = \begin{pmatrix} [f'(x^*)]_{\mathcal{B}} u_{\mathcal{B}} \\ 0 \end{pmatrix},$$

and since we have assumed that $h'(x^*, y^*)$ is nonsingular, we must have $[f'(x^*)]_{\mathcal{B}} u_{\mathcal{B}} \neq 0$ for $u_{\mathcal{B}} \neq 0$. Hence, (6.1) is nonsingular.

We now assume that (6.1) is nonsingular and that $x^* + f(x^*) > 0$, and we prove that $h'(x^*, y^*)$ is nonsingular for $y^* = f(x^*)$. If

$$h'(x^*, y^*) \begin{pmatrix} u \\ v \end{pmatrix} = 0,$$

for some vectors u and v, then the definition of h' implies that

$$f'(x^*)u = v, \qquad y_i^* u_i + x_i^* v_i = 0.$$

If $i \in \mathcal{B}$, then $y_i^* = 0$ and $x_i^* > 0$, and thus $v_i = 0$. Moreover, if $i \notin \mathcal{B}$, then $y_i^* > 0$ and $x_i^* = 0$, so that $u_i = 0$. Hence, $f'(x^*)u = v$ implies that $[f'(x^*)]_{\mathcal{B}} u_{\mathcal{B}} = 0$, and since we assumed nonsingularity of this submatrix, $u_{\mathcal{B}} = 0$. We have shown that $u = 0$, and thus $v = f'(x^*)u = 0$. This proves that $h'(x^*, y^*)$ is nonsingular. ■

We want to use Theorem 6.1 to prove that if $\{(x_k, y_k)\}$ is the sequence generated by the trust region Newton method, and if x^* is a limit point of $\{x_k\}$ that satisfies the assumptions of Theorem 6.1, then the whole sequence converges to x^*.

Theorem 6.2 *Let $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ be continuously differentiable on $\mathbb{R}^n_+$, and let $\{(x_k, y_k)\}$ be the sequence generated by the trust region Newton method. Assume that $\{x_k\}$ is bounded, and let x^* be a limit point of $\{x_k\}$. If x^* is a solution to the nonlinear complementarity problem* (1.1) *such that $x^* + f(x^*) > 0$ and* (6.1) *is nonsingular, then $\{(x_k, y_k)\}$ converges to (x^*, y^*) with $y^* = f(x^*)$.*

Proof. The first step in the proof is to show that z^* is a limit point of $\{z_k\}$. This result can be established by noting that since Theorem 5.1 guarantees that $\{h(z_k)\}$ converges to zero, $\{f(x_k) - y_k\}$ converges to zero. Thus, $f(x^*)$ is a limit point of $\{y_k\}$. This shows that z^* is a limit point of $\{z_k\}$.

For the rest of the proof we need to estimate the behavior of h near z^*. Theorem 6.1 guarantees that $h'(z^*)$ is nonsingular, and thus we can choose $\delta > 0$ so that $h'(z)$ is nonsingular for $\|z - z^*\| \leq \delta$. Let $\varepsilon > 0$ be such that

$$\varepsilon \|s_k\| \leq \|h'(z) s_k\| \tag{6.2}$$

for $\|z - z^*\| \le \delta$. Moreover, since $h(z^*) = 0$, we can also require that

$$\varepsilon \|z - z^*\| \le \|h(z)\| \tag{6.3}$$

whenever $\|z - z^*\| \le \delta$. Finally, choose an index k_0 such that $\|z_{k_0} - z^*\| \le \delta$ and

$$\|h(z_{k_0})\| \le \tfrac{1}{4}\varepsilon\delta. \tag{6.4}$$

This is possible because $\{h(z_k)\}$ converges to zero. We use an induction argument to show that these estimates imply that $\|z_k - z^*\| \le \delta$ for all $k \ge k_0$.

The first step in the induction argument is satisfied by the choice of the index k_0. Assume that $\|z_k - z^*\| \le \delta$ for some $k \ge k_0$, and note that (6.3) and (6.4) imply that

$$\varepsilon \|z_k - z^*\| \le \|h(z_k)\| \le \|h(z_{k_0})\| \le \tfrac{1}{4}\varepsilon\delta.$$

Hence, $\|z_k - z^*\| \le \frac{1}{4}\delta$. We bound the size of s_k by noting that since $\psi_k(s_k) \le 0$,

$$\|h(z_k) + h'(z_k)s_k\| \le \|h(z_k)\|,$$

and thus

$$\|h'(z_k)s_k\| \le 2\|h(z_k)\|.$$

This inequality, together with (6.2) and (6.4), implies that

$$\varepsilon \|s_k\| \le \|h'(z_k)s_k\| \le 2\|h(z_k)\| \le 2\|h(z_{k_0})\| \le \tfrac{1}{2}\varepsilon\delta.$$

Hence, $\|s_k\| \le \frac{1}{2}\delta$, and thus $\|z_{k+1} - z^*\| \le \delta$. This completes the induction argument and shows that $\|z_k - z^*\| \le \delta$ for all $k \ge k_0$. Convergence of $\{z_k\}$ to z^* is then a direct consequence of (6.3) and the convergence of $\{h(z_k)\}$ to zero. ■

The rate of convergence of the sequence generated by the trust region method depends on the choice of s_k. We base the computation of the step s_k on the subproblem

$$\min \{ q_k(z) : z_i = 0, \ i \in \mathcal{A}(z_{k,1}) \}, \tag{6.5}$$

where $z_{k,1} = z_k + s_k^C$, the quadratic $q_k : \mathbb{R}^n \mapsto \mathbb{R}$ is defined by

$$q_k(z) = \tfrac{1}{2} \left\| h(z_k) + h'(z_k)(z - z_k) \right\|^2,$$

and $\mathcal{A}(z_{k,1})$ is the active set at $z_{k,1}$, that is, the constraints i such that $[z_{k,1}]_i = 0$.

We can always choose $s_k = s_k^C$, but it is usually desirable to reduce q_k further. The subproblem (6.5) is an unconstrained linear least squares problem, so it is not difficult to compute a descent direction $p_{k,1}$ with $[p_{k,1}]_i = 0$ for $i \in \mathcal{A}(z_{k,1})$. We can then examine q_k in the ray $z_{k,1} + \beta p_{k,1}$, with $\beta \ge 0$, and choose $\beta_{k,1}$ so that q_k is minimized; if q_k does not have a minimum, choose $\beta_{k,1} = +\infty$. The minor iterate $z_{k,2} = z_{k,1} + \beta_{k,1} p_{k,1}$ may not be acceptable either because $z_{k,2}$ is not feasible or because $z_{k,2}$ does not satisfy the trust region constraint $\|z_{k,2} - z_k\| \le \Delta_k$. Thus, if necessary, we modify $\beta_{k,1}$ so that both constraints are satisfied.

The descent direction $p_{k,1}$ can be generated by either direct or iterative methods. The use of iterative methods is usually advised for large problems, while direct methods tend to be more reliable for small and medium-sized problems.

Instead of using a line search to determine $z_{k,2}$ we could use a projected search along the path defined by $[z_{k,1} + \beta p_{k,1}]_+$. The advantage of this approach is that we would be able to add several constraints at once. In this case we would not insist on strict decrease of q_k from $z_{k,1}$ to $z_{k,2}$ because this would require the determination of the first local minimizer. Instead we would require that the sufficient decrease condition

$$q_k(z_{k,2}) \le q_k(z_{k,1}) + \mu_0 \nabla q_k(z_{k,1})^T (z_{k,2} - z_{k,1})$$

be satisfied. This is precisely the same condition that we require for the computation of the Cauchy point. For additional details on projected searches, see Section 4 of Moré and Toraldo [36].

The process that we have outlined above can be repeated to generate a sequence of minor iterates $z_{k,1}, z_{k,2}, \ldots, z_{k,l}$. Global convergence is obtained as long as the quadratic decreases at each stage, but a superlinear rate of convergence requires a stronger requirement which is discussed later.

Calculation of step. Let $z_{k,0} = z_k$, and compute l minor iterates $z_{k,1}, z_{k,2}, \ldots, z_{k,l}$ with

$$z_{k,j} \in \mathbb{R}^{2n}_+, \qquad \|z_{k,j} - z_k\| \le \mu_1 \Delta_k, \qquad z_{k,1} = z_k + s_k^C,$$

and such that the sufficient decrease condition

$$q_k(z_{k,j+1}) \le q_k(z_{k,j}) + \mu_0 \nabla q_k(z_{k,j})^T (z_{k,j+1} - z_{k,j}) \tag{6.6}$$

is satisfied. The step is then defined by $s_k = z_{k,l} - z_k$.

We have assumed that we always compute a fixed number l of minor iterates $z_{k,j}$. This only imposes an upper bound on the number of minor iterates because we can always set $z_{k,j+1} = z_{k,j}$.

The conditions on the minor iterates are similar to those used by Lescrenier [27]. However, Lescrenier assumed that

$$q_k(z_{k,j+1}) \le q_k(\alpha z_{k,j+1} + (1-\alpha) z_{k,j}), \qquad \alpha \in [0,1]. \tag{6.7}$$

This requirement can be satisfied if a line search is used to choose the minor iterates, but it rules out the projected searches that we have proposed. Also note that our convergence analysis does not require Lescrenier's [27] assumption that

$$\mathcal{A}(z_k + s_k^C) \subset \mathcal{A}(z_k + s_k).$$

This condition rules out, for example, choosing $z_k + s_k$ in the interior of $\mathbb{R}^n_+$.

An important observation is that assumption (6.7) on the minor iterates is stronger than (6.6) when q_k is a quadratic. This observation can be verified by proving that if $\phi : \mathbb{R} \mapsto \mathbb{R}$ is a convex quadratic on $[0, 1]$, and

$$\phi(1) \le \phi(\alpha), \qquad \alpha \in [0, 1],$$

then

$$\phi(1) \le \phi(0) + \tfrac{1}{2}\phi'(0).$$

Indeed, since $\phi(1) \le \phi(\alpha)$ for $\alpha \in [0, 1]$, we must have $\phi'(1) \le 0$, and since ϕ is a quadratic,

$$2\left(\phi(1) - \phi(0)\right) = \phi'(1) + \phi'(0).$$

The result is now a direct consequence of these remarks.

The rate of convergence results depend on showing that eventually the trust region bound is not active. This result requires an estimate of the decrease of the quadratic q_k. The estimates needed for this result are valid for a general strictly convex function.

Lemma 6.3 *Assume that $\phi : \mathbb{R} \mapsto \mathbb{R}$ is twice differentiable on $[0, 1]$ and that there is a $\varepsilon > 0$ such that $\phi''(\alpha) \ge \varepsilon$ on $[0, 1]$. If*

$$\phi(1) \le \phi(0) + \mu\phi'(0) \tag{6.8}$$

for some $\mu \in (0, \frac{1}{2})$, then

$$\phi(0) - \phi(1) \ge \frac{\mu}{2(1-\mu)}\varepsilon.$$

Proof. The mean value theorem shows that

$$\phi(1) = \phi(0) + \phi'(0) + \tfrac{1}{2}\phi''(\theta)$$

for some $\theta \in (0, 1)$, and thus (6.8) implies that

$$\tfrac{1}{2}\phi''(\theta) \le (1-\mu)(-\phi'(0)).$$

Hence,

$$\phi(0) - \phi(1) \ge \mu(-\phi'(0)) \ge \frac{\mu}{2(1-\mu)}\phi''(\theta) \ge \frac{\mu}{2(1-\mu)}\varepsilon,$$

as desired. ■

This result has immediate application to the analysis of the trust region method. If we define

$$\phi(\alpha) = q_k\left(\alpha z_{k,j+1} + (1-\alpha) z_{k,j}\right)$$

and assume that the sequence $\{z_k\}$ converges to $z^* = (x^*, y^*)$ with $h'(x^*, y^*)$ nonsingular, then there is an $\varepsilon > 0$ such that

$$\varepsilon \|v\| \leq \|h'(z_k)v\|, \qquad v \in \mathbb{R}^{2n}$$

for all $k \geq 0$ sufficiently large. Hence, Lemma 6.3 implies that

$$q_k(z_{k,j}) - q_k(z_{k,j+1}) \geq \varepsilon \|z_{k,j+1} - z_{k,j}\|^2. \tag{6.9}$$

We will need this estimate for our next result.

Theorem 6.4 *Let $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ be continuously differentiable on $\mathbb{R}^n_+$ with f' Lipschitz continuous in a neighborhood of x^*, and let $\{(x_k, y_k)\}$ be the sequence generated by the trust region Newton method. Assume that $\{x_k\}$ converges to a solution x^* of the nonlinear complementarity problem* (1.1) *such that $x^* + f(x^*) > 0$ and* (6.1) *is nonsingular. If the step s_k is calculated as specified above, then the trust region bound Δ_k is bounded away from zero.*

Proof. In the proof we bound $|\rho_k - 1|$, where ρ_k is defined by (5.4), and show that the bounds converge to zero; the rules for updating Δ_k then show that Δ_k is bounded away from zero. We begin by noting that

$$\rho_k - 1 = \frac{\|h(z_k + s_k)\|^2 - \|h(z_k) + h'(z_k)s_k\|^2}{\psi_k(s_k)}.$$

The denominator of this expression is estimated by proving that there is an $\varepsilon_0 > 0$ such that

$$-\psi_k(s_k) \geq \varepsilon_0 \|s_k\|^2. \tag{6.10}$$

Since we have already established (6.9), we obtain that the decrease generated by s_k satisfies

$$q_k(z_k) - q_k(z_k + s_k) \geq \nu \sum_{j=1}^{l} \|z_{k,j} - z_{k,j-1}\|^2 \geq \varepsilon \max_{1 \leq j \leq l} \left\{ \|z_{k,j} - z_{k,j-1}\|^2 \right\}.$$

On the other hand,

$$\|s_k\| \leq \sum_{j=1}^{l} \|z_{k,j} - z_{k,j-1}\| \leq l \max_{1 \leq j \leq l} \{\|z_{k,j} - z_{k,j-1}\|\}.$$

Hence, (6.10) holds with $\varepsilon_0 = \varepsilon / l$. We estimate the numerator of the expression for $\rho_k - 1$ by

$$\left| \|h(z_k + s_k)\|^2 - \|h(z_k) + h'(z_k)s_k\|^2 \right| \leq \mu_k \left\| h(z_k + s_k) - h(z_k) - h'(z_k)s_k \right\|,$$

where

$$\mu_k = \|h(z_k + s_k)\| + \|h(z_k) + h'(z_k)s_k\|,$$

and by

$$\|h(z_k+s_k)-h(z_k)-h'(z_k)s_k\| \le \sup_{0\le\theta\le 1}\left\{\|h'(z_k+\theta s_k)-h'(z_k)\|\right\}\|s_k\| \le \kappa\|s_k\|^2,$$

where κ is the Lipschitz constant for h'. For this last estimate we made use of the fact that the Lipschitz continuity of f' implies that h' is also Lipschitz continuous near x^*.

These estimates show that $|\rho_k-1| \le (\kappa/\varepsilon_0)\mu_k$, so that the our result will be established if we show that $\{\mu_k\}$ converges to zero. Since $\{z_k\}$ converges to z^* with $h(z^*)=0$, we obtain that $\{\mu_k\}$ converges to zero if $\{s_k\}$ converges to zero. Note that $\psi_k(s_k)\le 0$, and thus

$$\|h(z_k)+h'(z_k)s_k\| \le \|h(z_k)\|.$$

Hence, the nonsingularity of $h'(x^*,y^*)$ implies that

$$\varepsilon\|s_k\| \le \|h'(z_k)s_k\| \le 2\|h(z_k)\|.$$

This estimate clearly shows that $\{s_k\}$ converges to zero. $\blacksquare$

Theorem 6.4 requires that $z_{k,j}$ satisfy (6.6), and thus the Cauchy step s_k^C is acceptable. A superlinear rate of convergence requires that we impose further conditions on s_k.

When the iterate z_k is far away from the solution, the step s_k is usually determined because the trust region bound $\|z_{k,j}-z_k\| \le \mu_1\Delta_k$ is encountered during the computation of $z_{k,j+1}$. However, as we converge, Theorem 6.4 shows that the trust region does not interfere with the computation of the step, so that we are free to reduce q_k further by searching the feasible set. We propose to continue computing minor iterates until $z_{k,l}$ is an approximate minimizer of q_k on the current active set $\mathcal{A}(z_{k,l})$. If P_k is the projection into the subspace

$$\left\{z\in\mathbb{R}^{2n} : z_i=0,\ i\in\mathcal{A}(z_{k,l})\right\},$$

then we require that

$$\left\|Q_k^T[h(z_k)+h'(z_k)s_k]\right\| \le \xi_k\,\|h(z_k)\|, \tag{6.11}$$

where Q_k is the Jacobian matrix with respect to the free variables, that is,

$$Q_k = h'(z_k)P_k.$$

We can motivate this requirement by noting that if $\Psi_k(w)=q_k\left(z_{k,l}+P_kw\right)$, then

$$\nabla\Psi_k(0) = Q_k^T[h(z_k)+h'(z_k)s_k].$$

In particular, if we choose $\xi_k=0$, then $z_{k,l}$ is a minimizer with respect to $\mathcal{A}(z_{k,l})$.

We have already noted that the step s_k is usually determined because the trust region bound $\|z_{k,j}-z_k\| \le \mu_1\Delta_k$ is encountered during the computation of $z_{k,j+1}$. Thus, we only need to assume that the step s_k satisfies (6.11) if $\|s_k\| \le \mu_*\Delta_k$ for some $\mu_* < \mu_1$.

Theorem 6.5 *Let $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ be continuously differentiable on $\mathbb{R}^n_+$ with f' Lipschitz continuous in a neighborhood of x^*, and let $\{(x_k, y_k)\}$ be the sequence generated by the trust region Newton method. Assume that $\{x_k\}$ converges to a solution x^* of the nonlinear complementarity problem* (1.1) *such that $x^* + f(x^*) > 0$ and* (6.1) *is nonsingular. If the step s_k is calculated by the algorithm outlined above, and* (6.11) *holds whenever $\|s_k\| \le \mu_* \Delta_k$ for some $\mu_* < \mu_1$, then $\{z_k\}$ converges Q-linearly to z^* when ξ^* is sufficiently small, where*

$$\xi^* = \limsup_{k \to +\infty} \xi_k.$$

The rate of convergence is Q-superlinear when $\xi^* = 0$, and Q-quadratic when

$$\xi_k \le \sigma \|h(z_k)\|$$

for some constant $\sigma > 0$.

Proof. The main estimate that we need for this result is that

$$\varepsilon \, \|z_{k+1} - z^*\| \le \left\| Q_k^T h(z_{k+1}) \right\| + \varepsilon_k \, \|z_{k+1} - z^*\| \tag{6.12}$$

for some $\varepsilon > 0$ and some sequence $\{\varepsilon_k\}$ converging to zero. Note that

$$Q_k^T h(z_{k+1}) = Q_k^T h'(z_k)(z_{k+1} - z^*) + Q_k^T [h(z_{k+1}) - h(z^*) - h'(z_k)(z_{k+1} - z^*)],$$

and thus standard estimates of the last term show that

$$\left\| Q_k^T h'(z_k)(z_{k+1} - z^*) \right\| \le \left\| Q_k^T h(z_{k+1}) \right\| + \varepsilon_k \|z_{k+1} - z^*\|,$$

where $\{\varepsilon_k\}$ converges to zero. We now note that

$$P_k(z_{k+1} - z^*) = z_{k+1} - z^*,$$

since $\mathcal{A}(z_{k+1}) \subset \mathcal{A}(z^*)$, and thus

$$\left\| Q_k^T h'(z_k)(z_{k+1} - z^*) \right\| = \left\| Q_k^T Q_k (z_{k+1} - z^*) \right\| \ge \varepsilon \, \|z_{k+1} - z^*\|,$$

where ε is a lower bound on the eigenvalues of $h'(z_k)^T h'(z_k)$. Hence, (6.12) holds.

The result follows from (6.12) by standard arguments. Since

$$\left\| Q_k^T h(z_{k+1}) \right\| \le \left\| Q_k^T [h(z_{k+1}) - h(z_k) - h'(z_k) s_k] \right\| + \left\| Q_k^T [h(z_k) + h'(z_k) s_k] \right\|,$$

we obtain that

$$\left\| Q_k^T h(z_{k+1}) \right\| \le \kappa \|s_k\|^2 + \xi_k \, \|h(z_k)\|,$$

where κ is the Lipschitz constant for h'. We also know that there are constants μ_1 and μ_2 such that

$$\|s_k\| \le \mu_1 \|h(z_k)\|, \qquad \|h(z_k)\| \le \mu_2 \|z_k - z^*\|.$$

The first inequality was established at the end of the proof of Theorem 6.4, while the second inequality follows from the existence of $h'(z^*)$ with $h(z^*) = 0$. Hence, (6.12) shows that

$$\varepsilon \|z_{k+1} - z^*\| \leq \kappa(\mu_1\mu_2)^2 \|z_k - z^*\|^2 + \mu_2\xi_k \|z_k - z^*\| + \varepsilon_k \|z_{k+1} - z^*\| . \tag{6.13}$$

This inequality shows that $\{z_k\}$ converges Q-linearly to z^* if $\mu_2\xi^* < \varepsilon$. Inequality (6.13) also shows that the rate of convergence is Q-superlinear when $\xi^* = 0$, and Q-quadratic when $\xi_k \leq \sigma\|h(z_k)\|$ for some constant $\sigma > 0$. ■

A weakness in our convergence analysis is that we are not able to prove Q-linear convergence for any $\xi^* < 1$. This may not be possible since for degenerate problems the active set $\mathcal{A}(z_k)$ may not settle down; R-linear convergence seems to be possible.

The convergence results of Pang and Gabriel [39], and Gabriel and Pang [14], require that the limit point x^* be s-regular and b-regular. These regularity assumptions imply that (6.1) is nonsingular, but do not require our assumption that $x^* + f(x^*) > 0$.

The algorithms proposed by Pang and Gabriel [39], Gabriel and Pang [14, 15], and Monteiro, Pang, and Wang [32] are quite different from the trust region method. From a computational viewpoint, an important difference is that each iteration of these algorithms requires the solution of a linear programming problem or a quadratic programming problem, while the trust region method requires the solution of systems of linear equations.

Acknowledgments. A special thanks goes to Stephen Wright for his advice and suggestions during the preparation of this work. I am also grateful to Michael Ferris, Steven Gabriel, and Danny Ralph for their remarks on the final version of this manuscript. Gail Pieper's gentle comments are appreciated, as well as her attempts to explain subtle differences in the meaning of words.

References

[1] B. H. Ahn, *Computation of Market Equilibria for Policy Analysis: The Project Independence Evaluation Study (PIES) Approach*, Garland, 1979.

[2] D. P. Bertsekas, *Projected Newton methods for optimization problems with simple constraints*, SIAM J. Control Optim., 20 (1982), pp. 221–246.

[3] J. V. Burke, J. J. Moré, and G. Toraldo, *Convergence properties of trust region methods for linear and convex constraints*, Math. Programming, 47 (1990), pp. 305–336.

[4] B. Chen and P. T. Harker, *A continuation method for monotone variational inequalities*, Preprint 90-10-02, University of Pennsylvania, Philadelphia, Pennsylvania, 1990. Revised April 1993.

[5] T. F. Coleman and Y. Li, *An interior trust region approach for nonlinear minimization subject to bounds*, Technical report 93–1342, Cornell University, Ithaca, New York, 1993.

[6] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Global convergence of a class of trust region algorithms for optimization problems with simple bounds*, SIAM J. Numer. Anal., 25 (1988), pp. 433–460.

[7] ——, *Testing a class of methods for solving minimization problems with simple bounds on the variables*, Math. Comp., 50 (1988), pp. 399–430.

[8] R. W. Cottle, J. S. Pang, and R. E. Stone, *The Linear Complementarity Problem*, Academic Press, 1992.

[9] S. P. Dirkse and M. C. Ferris, *The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems*, Report 1179, University of Wisconsin, Madison, Wisconsin, 1993.

[10] ——, *MCPLIB: A collection of nonlinear mixed complementarity problems*, Report 1215, University of Wisconsin, Madison, Wisconsin, 1994.

[11] M. Fiedler and V. Pták, *On matrices with non-positive off-diagonal elements and positive principal minors*, Czech. Math. J., 12 (1962), pp. 382–400.

[12] ——, *Some generalizations of positive definiteness and monotonicity*, Numer. Math., 9 (1966), pp. 163–172.

[13] M. Fukushima, *Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems*, Math. Programming, 53 (1992), pp. 99–110.

[14] S. A. Gabriel and J. S. Pang, *An inexact NE/SQP method for solving the nonlinear complementarity problem*, Comp. Opt. Appl., 1 (1992), pp. 67–91.

[15] ——, *A trust region method for constrained nonsmooth equations*, in Large-Scale Optimization: State of the Art, W. W. Hager, D. W. Hearn, and P. M. Pardalos, eds., Kluwer Academic Publishers, Inc., 1993, pp. 000–000.

[16] D. M. Gay, *A trust region approach to linearly constrained optimization*, in Numerical Analysis, D. F. Griffiths, ed., Lecture Notes in Mathematics 1066, Springer-Verlag, 1984, pp. 72–105.

[17] C. Geiger and C. Kanzow, *On the resolution of monotone complementarity problems*, Preprint 82, Universität Hamburg, Hamburg, Germany, 1994.

[18] P. T. Harker and J. S. Pang, *Finite dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications*, Math. Programming, 48 (1990), pp. 161–220.

[19] P. T. Harker and B. Xiao, *Newton's method for the nonlinear complementarity problem: A B-differentiable equation problem*, Math. Programming, 48 (1990), pp. 339–357.

[20] C. Kanzow, *Nonlinear complementarity as unconstrained optimization*, Preprint 67, Universität Hamburg, Hamburg, Germany, 1993.

[21] ——, *Some equation-based methods for the nonlinear complementarity problem*, Preprint 63, Universität Hamburg, Hamburg, Germany, 1993.

[22] M. Kojima, N. Megiddo, and S. Mizuno, *A general framework of continuation method for complementarity problems*, Math. Oper. Res., 18 (1993), pp. 945–963.

[23] M. Kojima, N. Megiddo, and T. Noma, *Homotopy continuation methods for nonlinear complementarity problems*, Math. Oper. Res., 16 (1991), pp. 754–774.

[24] M. Kojima, S. Mizuno, and T. Noma, *A new continuation method for complementarity problems with uniform P-functions*, Math. Programming, 43 (1989), pp. 107–113.

[25] ——, *Limiting behavior of trajectories generated by a continuation method for monotone complementarity problems*, Math. Oper. Res., 15 (1990), pp. 662–675.

[26] M. Kojima, T. Noma, and A. Yoshise, *Global convergence and detecting infeasibility in interior point algorithms*, Report B-257, Tokyo Institute of Technology, Tokyo, Japan, 1992.

[27] M. Lescrenier, *Convergence of trust region algorithms for optimization with bounds when strict complementarity does not hold*, SIAM J. Numer. Anal., 28 (1991), pp. 476–495.

[28] O. L. Mangasarian, *Equivalence of the complementarity problem to a system of nonlinear equations*, SIAM J. Appl. Math., 31 (1976), pp. 89–92.

[29] O. L. Mangasarian and M. V. Solodov, *Nonlinear complementarity as unconstrained and constrained minimization*, Math. Programming, 62 (1993), pp. 277–297.

[30] A. S. Manne, ed., *Economic Equilibrium: Model Formulation and Solution*, Mathematical Programming Study 23, North-Holland, 1985.

[31] L. McLinden, *The complementarity problem for maximal monotone multifunctions*, in Variational Inequalities and Complementarity Problems, R. W. Cottle, F. Giannessi, and J.-L. Lions, eds., John Wiley & Sons, 1980, pp. 251–268.

[32] R. D. C. Monteiro, J. S. Pang, and T. Wang, *A positive algorithm for the nonlinear complementarity problem*, preprint, Georgia Institute of Technology, Atlanta, Georgia, 1993.

[33] J. J. Moré, *Classes of function and feasibility conditions in nonlinear complementarity problems*, Math. Programming, 6 (1974), pp. 327–338.

[34] ——, *Coercivity conditions in nonlinear complementarity problems*, SIAM Rev., 16 (1974), pp. 1–16.

[35] J. J. Moré and W. C. Rheinboldt, *On P- and S-functions and related classes of n-dimensional mappings*, Linear Algebra Appl., 6 (1973), pp. 45–68.

[36] J. J. Moré and G. Toraldo, *On the solution of large quadratic programming problems with bound constraints*, SIAM J. Optimization, 1 (1991), pp. 93–113.

[37] A. Nagurney, ed., *Advances in Equilibrium Modeling, Analysis and Computation*, Annals of Operations Research 44, J. C. Baltzer AG, 1993.

[38] J. S. Pang, *Newton's method for B-differentiable equations*, Math. Oper. Res., 15 (1990), pp. 311–341.

[39] J. S. Pang and S. A. Gabriel, *NE/SQP: A robust algorithm for the nonlinear complementarity problem*, Math. Programming, 60 (1993), pp. 295–337.

[40] D. Ralph, *Global convergence of damped Newton's method for nonsmooth equations via the path search*, Math. Oper. Res., 19 (1994), pp. 000 – 000.

[41] S. M. Robinson, *Normal maps induced by linear transformations*, Math. Oper. Res., 17 (1992), pp. 691–714.

[42] J. Soares and J. Júdice, *An active set Newton's method for large-scale nonlinear programs with box constraints*, preprint, Universidada de Coimbra, 1993.

[43] P. K. Subramanian, *Gauss-Newton methods for complementarity problems*, J. Optim. Theory Appl., 77 (1993), pp. 467–482.

[44] L. T. Watson, *Solving the nonlinear complementarity problem by a homotopy method*, SIAM J. Control Optim., 17 (1976), pp. 36–46.

[45] S. J. Wright, *A superlinear infeasible interior-point algorithm for monotone complementarity problems*, Preprint MCS-P344-1292, Argonne National Laboratory, Argonne, Illinois, 1992. Revised July 1993.

[46] ——, *Stability of linear equations solvers in interior-point methods*, Preprint MCS-P400-1293, Argonne National Laboratory, Argonne, Illinois, 1993.

[47] B. Xiao and P. T. Harker, *A nonsmooth Newton method for variational inequalities, I: Theory*, preprint, Seton Hall University, South Orange, New Jersey, 1993.

[48] ——, *A nonsmooth Newton method for variational inequalities, II: Numerical results*, preprint, Seton Hall University, South Orange, New Jersey, 1993.